Nature Conservation 34: 477–503 (2019)
doi: 10.3897/natureconservation.34.30735
http://natureconservation.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
Nature Conservation
Launched to accelerate biodiversity conservation

Feeding Essential Biodiversity Variables (EBVs): actual and potential contributions from LTER-Italy

Martina Zilioli[1], Alessandro Oggioni[1], Paolo Tagliolato[1], Alessandra Pugnetti[2], Paola Carrara[1]

1 *National Research Council of Italy, Institute for Electromagnetic Sensing of the Environment, via Corti 12 - 20133 Milano, Italy* **2** *National Research Council of Italy, Institute of Marine Sciences, Arsenale Tese 104, Castello 2737/F, 30122, Venezia, Italy*

Corresponding author: *Martina Zilioli* (zilioli.m@irea.cnr.it)

Academic editor: *M. Freppaz* | Received 21 October 2018 | Accepted 8 March 2019 | Published 3 May 2019

http://zoobank.org/81D46C70-1FAC-4557-8F0C-B084907018C9

Citation: Zilioli M, Oggioni A, Tagliolato P, Pugnetti A, Carrara P (2019) Feeding Essential Biodiversity Variables (EBVs): actual and potential contributions from LTER-Italy. In: Mazzocchi MG, Capotondi L, Freppaz M, Lugliè A, Campanaro A (Eds) Italian Long-Term Ecological Research for understanding ecosystem diversity and functioning. Case studies from aquatic, terrestrial and transitional domains. Nature Conservation 34: 477–503. https://doi.org/10.3897/natureconservation.34.30735

Abstract

The conceptual framework of Essential Biodiversity Variables (EBVs) aims to capture the major dimensions of biodiversity change by structuring biodiversity monitoring and by ruling data collection amongst different providers. Amongst the research infrastructures adopting and implementing the EBV framework, LTER-Europe - the European node of ILTER (International Long-Term Ecological Research) - follows the approach to compare site-based biodiversity observations within and across its networks. However, a synoptic overview of their contributions with EBVs-relevant data is still missing, since data are not made available for several reasons. In this paper, we assess the capacity of LTER-Italy, one of the richest and heterogeneous networks of LTER sites in Europe, to provide data to "Species Distribution" and "Species Abundance" EBVs without inspecting and downloading their contents. To this aim, we mine the EBVs information which is publicly structured and shared by LTER site managers through DEIMS-SDR, the LTER-Europe online metadata repository. We classify the sites according to two types of contributions: (i) the *actual* contribution, based on metadata of datasets and (ii) the *potential* contribution, based on metadata of sites. Through these assessments, we investigate if LTER-Italy monitoring activities can provide EBVs measures and which sites currently provide datasets. By comparing the two contributions,

Copyright *Martina Zilioli et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

we pinpoint the factors hampering the accessibility of LTER-Italy data and suggest solutions to increase the discoverability and reusability of LTER-Italy EBVs measurements. The research provides the first overview of EBVs monitored in LTER-Italy and the corresponding data management practices, as well as an evaluation of the interoperability of this network with respect to other research organisations for legal and technical aspects.

Keywords

Essential Biodiversity Variables, LTER-Italy, DEIMS-SDR, metadata analysis, research infrastructure assessment, EML, EMF

Introduction

Despite its indisputable role for human well-being and for ecosystem functioning (Díaz et al. 2006), biodiversity is threatened by anthropogenic stressors (Barnosky et al. 2011; Dirzo et al. 2014). The United Nations Convention on Biological Diversity (CBD) was organised to encourage countries to reduce pressures on biodiversity. In 2010, the CBD identified the "Aichi Targets" as specific goals for 2011–2020, to steer subscribing parties in developing national plans for assessing biodiversity loss and for providing solutions through policy regulations. Despite the increasing data volume in the last decades (Kelling et al. 2009), the fact that two-thirds of reports previously submitted to CBD lacked evidence-based information on biodiversity change (Bubb and Chenery 2011; UNEP CBD 2010) revealed the need for better managing this knowledge to actually enable meaningful estimation for policy practices.

Proliferation of studies is not always accompanied by integration of data in decision-making (Sutherland et al. 2012) as the exchange of knowledge between science and policy requires brokering and it is slowed by the lack of information effectively asked by policy-makers (Sutherland et al. 2012) or by the lack of access to reliable datasets through adequate tools.

The conceptual framework of Essential Biodiversity Variables (EBVs) was endorsed by the CBD (UNEP CBD 2010, 2016) to address this issue, wherein it defines a minimum set of essential measurements to facilitate the reporting of data amongst practitioners (Pereira et al. 2013) and to quantify the major dimensions of biodiversity change. This is supported by different works (Pereira et al. 2013; Weltzin et al. 2014; Geijzendorffer et al. 2016), where the EBVs framework is used as an abstraction layer of measurements by means of which the primary observations captured from any biodiversity initiative can be related to Aichi Targets. The Group on Earth Observations Biodiversity Observation Network (GEO BON) developed the framework and grouped 22 EBVs in six main classes (i.e. Genetic Composition, GC; Species Populations, SP; Species Traits, ST; Community Composition, CC; Ecosystem Structure, ES; Ecosystem Function, EF), each representing a level of biodiversity organisation, which requires appropriate datasets. The comprehensive nature of this conceptual framework enables providers at any scale of study to help cover the six levels and allows prioritising data mobilisation if essential measurements are lacking in monitoring

programmes (Geijzendorffer et al. 2016). Both scholarly and citizens science projects (eBird; TEAM), as well as online publishers (Pangaea; GenBank; LPI) can provide EBVs-relevant data, while projects (e.g. GLOBIS-B) or worldwide observation systems such as GEO BON and the Global Ocean Observing System (Muller-Karger et al. 2018) currently work to make the framework operational at research level. Amongst these, the European node of the International Long Term Ecological Research network (LTER-Europe) is embracing the EBVs framework to compare biodiversity observations across its networks and sites (Haase et al. 2018; Mollenhauer et al. 2018). Within LTER-Europe, monitoring activities on multiple biotic and abiotic ecosystem parameters are carried out by research sites grouped in national scale observing systems, i.e. networks. According to the INSPIRE Thematic Working Group "Environmental Monitoring Facilities" (2012) (Wohner et al. 2018), the research sites are the environmental monitoring facilities in which the LTER networks are organised and those focusing on biodiversity are unique data sources for three reasons: (i) they provide biotic in-situ data "enriched" by complementary abiotic measures, (ii) they provide data with high temporal coverage as their activities are planned with long-term view and (iii) as they belong to a network system and are distributed in different places of a specific country, they provide data with high geographic coverage. As previously highlighted by other authors (Kissling et al. 2018a; Haase et al. 2018), LTER networks are good candidates to supply values for almost all EBVs classes by monitoring the three LTER realms (marine, terrestrial and freshwater) (Geijzendorffer et al. 2016). Additionally, long-term datasets are recommended data sources for CBD indicators (UNEP CBD 2010) and the LTER geographical organisation makes it possible to set-up a distributed environmental facility to provide measurements with specific country coverage.

Although ongoing efforts are undertaken to align monitoring programmes to the EBV concept, the capacity of LTER networks to deliver relevant data has not been described yet, even if reported by authors of EBVs studies (Geijzendorffer et al. 2016) and by scientific products of networks activities. The presence of such description would instead be very useful for governments and advisory bodies to be aware of the LTER role in collecting updated measurements and for biodiversity practitioners to assemble its data resources within a worldwide extent (Peterson et al. 2018; Schmeller et al. 2017; Hardisty et al. 2019a).

Integration of biodiversity datasets from multiple sources is one of the current challenges faced by ecological informatics. It requires the use of standardised measurement protocols, the adoption of common data standards, ontologies, the creation of controlled vocabularies (Rosati et al. 2017), the use of virtual laboratories (Hardisty et al. 2013) and tools for EBVs data visualisation as well (e.g. European Biodiversity Portal, EBP; Global Biodiversity Information Facility, GBIF). Recently, Kissling et al. (2018a) proposed the concept of a global EBV data product built by integrating existing heterogeneous primary data, appropriately harmonised. As proposed for "Species traits" (Kissling et al. 2018b), "Species abundance" and "Species distribution" (Kissling et al. 2018a), a standard data processing workflow is necessary for the aggregation of primary data into global data products. According to the number of data management and

analysis procedures which are accomplished in the pipeline, three types of harmonised datasets are generated (i.e. EBV-usable, EBV-ready, derived and modelled EBV data).

At the same time, metadata compiled in standardised forms are fundamental for aggregation of biodiversity datasets. Metadata support different processes of data integration, by facilitating the discovery and the reuse of generated data to other scientists (Michener 2006). In order to assess the fitness for purpose of primary datasets, the EBVs metadata should provide indications about the three dimensions of the variable (space, time and taxonomy) and related attributes (extent, resolution and measurement units) (Kissling et al. 2018a). Particularly, this information can be shared through different standards (Wilkinson et al. 2016), developed to allow machine-to-machine interaction and to provide comprehensive information to understand and reuse datasets, including that related to content, context, quality, structure and accessibility (Michener et al. 1997). Even if two Biodiversity Research Infrastructures (BRIs) are successfully tested to build EBV data products, limits to aggregation still remain (Hardisty et al. 2019b).

The EBVs framework is a theory-driven and academic approach to biodiversity monitoring. On the one hand, it helps to attain consensus on what is essential to monitor and where to focus the limited financial resources to assure the assessment of biodiversity change (Vihervaara et al. 2017). On the other hand, it does not establish methods and instrumentations to allow integration of measures (Haase et al. 2018; Hardisty et al. 2019b), scarcely fostering the combination of data from multiple sources or the attitude to share them through public repositories so that many potential data resources remain hidden to end-users.

To be reliable, the above-mentioned description of LTER EBVs-relevant data has to pinpoint how the data can be integrated without missing the identification of all the potential sources of the research infrastructure considered. In fact, although facilities such as global IT aggregators (e.g. GBIF) or e-Science infrastructures (e.g. LifeWatch) increase the access to different users, scientists apply restrictions to data (e.g. commercial use) by limiting access and confidential sharing practices hamper the review of their contents. Moreover, the lack of funding for data curation and publishing activities limit their sharing through digital archives.

The objective of our study is to demonstrate the capacity of LTER-Italy to provide EBVs data through the analysis of its metadata resources, by considering that: (I) data (e.g. the dataset itself) has not always been published for several reasons; (II) not all LTER sites measure the biodiversity components, but monitoring occurs according to the ecological research focus of the programme.

Materials and methods

To free the analysis of the LTER network from data inspection and to identify the specific causes of restricted access to datasets, we examined EBVs information structured in metadata of LTER sites and datasets published by site managers in the Dynamic Eco-

logical Information Management System – Site and Data Set Registry (DEIMS-SDR), that is the most comprehensive catalogue of field observations sites in environmental research networks (Mirtl et al. 2018; Wohner et al. 2019). We focused on LTER-Italy, which we introduce in the "Case study" subsection. In the "Mapping EBV information" subsection, we provide the first mapping of the EBV information in two DEIMS Metadata Models and in "Collection of EBV information from metadata" and the following subsections, we propose a method to collect and analyse the metadata compiled for LTER-Italy sites. As meaningful metrics, we define two types of contributions to "Species Abundance" (SA) and "Species Distribution" (SD) EBVs which belong to the "Species Populations" class: the potential contribution of the network, assessed by processing its Sites Metadata and the actual contribution of the network, based on its Data Set Metadata. While the former reveals the rate of sites that can provide EBV primary data, the latter reveals the rate of sites that are currently sharing metadata for SA and SD datasets. An overview of the workflow for obtaining the two metrics from site metadata and from datasets metadata, respectively, is depicted in Figure 1.

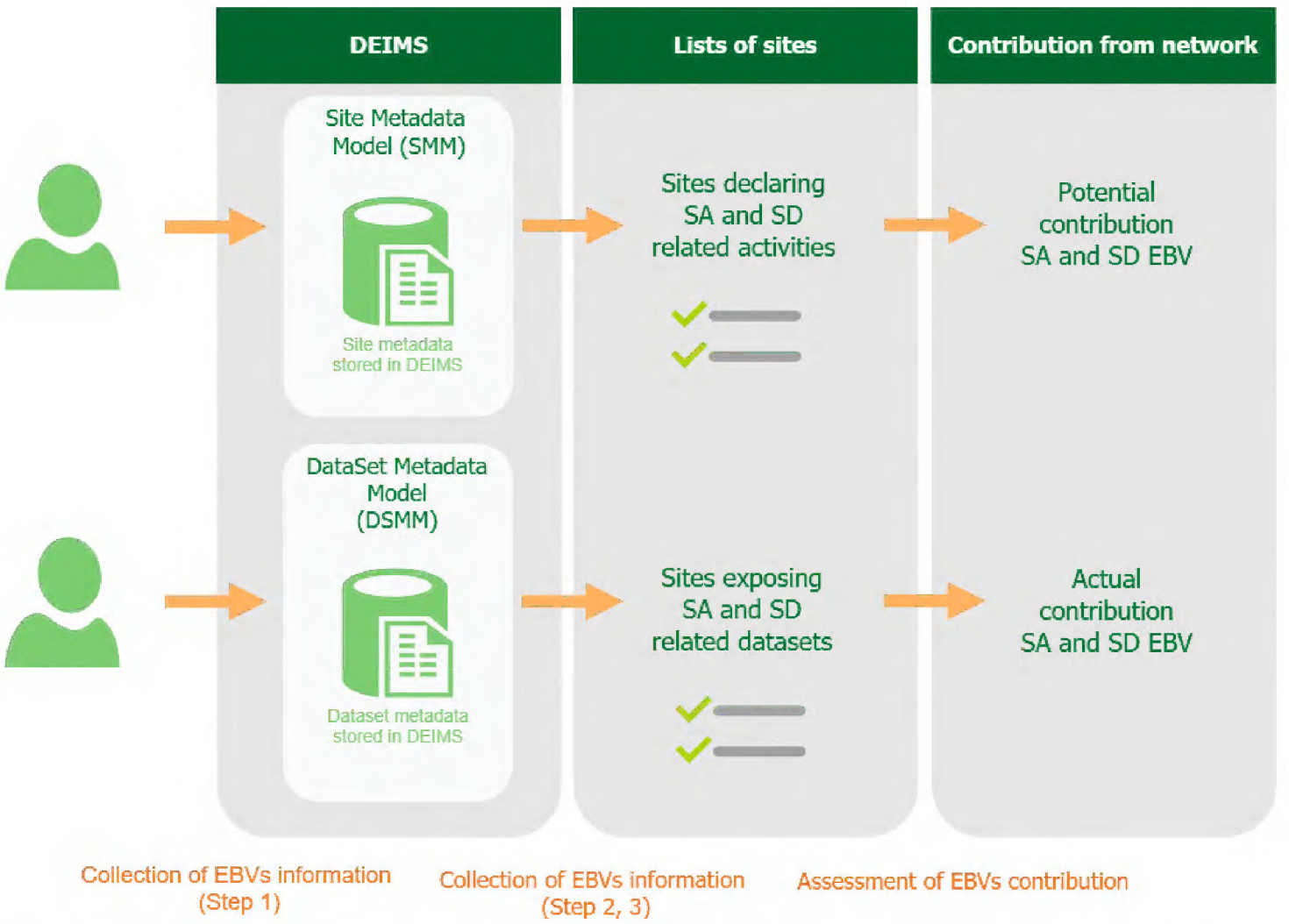
DEIMS
Lists of sites
Contribution from network
Site Metadata Model (SMM)
Site metadata stored in DEIMS
Sites declaring SA and SD related activities
Potential contribution SA and SD EBV
DataSet Metadata Model (DSMM)
Dataset metadata stored in DEIMS
Sites exposing SA and SD related datasets
Actual contribution SA and SD EBV
Collection of EBVs information (Step 1)
Collection of EBVs information (Step 2, 3)
Assessment of EBVs contribution

Figure 1. Metadata analysis overview. The diagram illustrates the activities required to perform the metadata analysis. The collection of EBV information from metadata is articulated in three steps which are followed by EBVs actual and potential contribution assessment for Species Abundance (abridged as "SA") and Species Distribution (abridged as "SD").

Case study

The LTER-Italy network is the Italian node of LTER-Europe and consists of 104 sites registered on DEIMS-SDR. It is the richest, amongst the European national LTER networks, with respect to the number of sites and it is one of the more heterogeneous for monitored ecosystems (Mollenhauer et al. 2018). In our analysis, we considered only those sites that are distributed inside the boundaries of the country, since our ultimate goal is to evaluate how this national monitoring programme can contribute to EBVs measurement or provide evidence-based data for local, regional and national governments in reporting to CBD; therefore, we excluded the 9 extraterritorial sites from the metadata analysis.

To avoid redundancy, we also excluded from our analysis the metadata from 23 Italian macrosites, as every macrosite aggregates the metadata of the sites it groups, which are individually analysed.

Hence, we analysed the metadata related to 72 sites and, in particular, we selected only those which compiled the metadata element "eLTER Parameter" (illustrated in the following subsection) which constitutes an informative tagging of the research activities of a site. The resulting set of sites is our statistical data sample and amounts to 43 sites.

Mapping EBV information

For the purpose of the present study, metadata of datasets and sites in LTER-Italy, stored in DEIMS-SDR, are queried: The two metadata models (DEIMS-SDR Metadata Models), in which these metadata are structured, are the Site Metadata Model (SMM) and the Data Set Metadata Model (DSMM). Both models contain elements referable to EBVs that allow us to assess whether a site can be an EBVs data provider and if a dataset can be reused (e.g. for its aggregation with other EBVs measurements). While the first model contains explicit references to EBVs, for the second we had to establish which elements should have been taken into consideration in our analysis. To this aim, we followed the metadata requirements described by Kissling et al. (2018a) and we checked which elements were actually compiled by site managers with this information.

EBVs information can be explicitly found from SMM in the "eLTER Parameters" element, whose content is a list of keywords from a hierarchically structured controlled vocabulary. The vocabulary is related to the LTER framework for standard observations (Haase et al. 2018; Mollenhauer et al. 2018) that combines the Ecosystem Integrity (EI) and the EBVs conceptual frameworks. Figure 2 illustrates the organisation of keywords relative to "Species Abundance" and "Species Distribution" observations.

As shown in Figure 2, the information relative to the EBVs monitoring programme is under the "Biodiversity" node of the tree structure of the vocabulary, further subdivided into three branches corresponding to the realms in which the EBVs are monitored: Terrestrial, Rivers and Lakes, Marine. These realms correspond to the biomes that

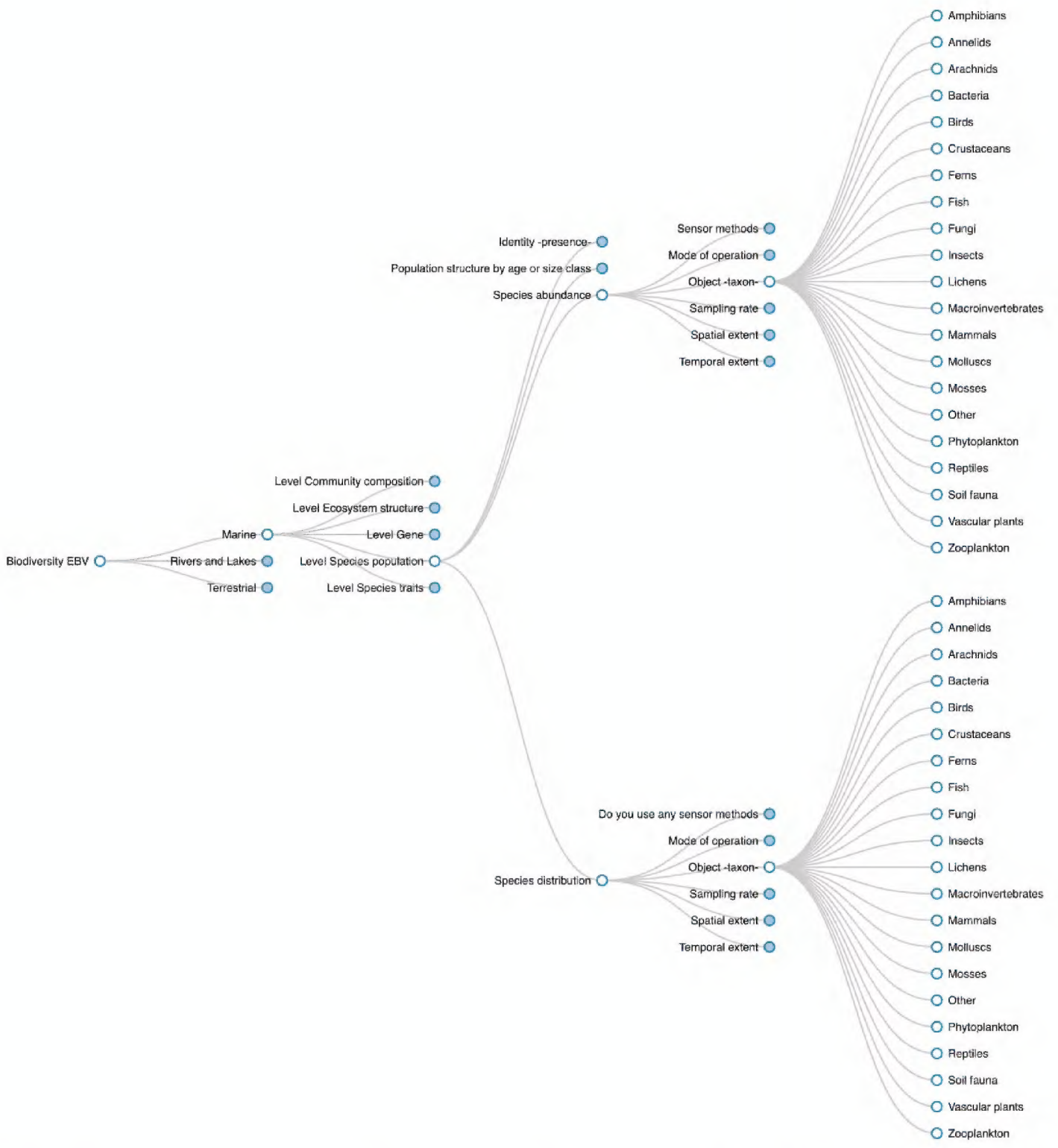

Figure 2. Organisation of EBVs-related keywords for the metadata element "eLTER Parameter". As an example, the Figure illustrates the tree structure for the Marine realm. The Figure shows the metadata field "Object (taxon)", associated with the eLTER Parameters element and analysed for LTER-Italy metadata. According to the realm selected, specific taxonomic terms are exposed. Empty circles provide the branches illustrated. Light blue circles are not expanded in the Figure.

are used for the classification of LTER-Italy sites in the following analysis. Each branch contains the following six fields related to aspects of the monitoring programme:

1. *Object (taxon)*, which represents information on the taxonomic extension of the measurements (in a site metadata document, the presence of a keyword from this branch indicates a taxonomic group of interest for the monitoring campaigns of the site);

2. *Spatial extent*, which corresponds to the spatial dimension of measurements, e.g. single sampling point, sampling surface;
3. *Temporal extent (sampling history)*, which corresponds to the temporal range of measurements;
4. *Mode of operation*, which describes the measurements process: either a continuous or campaign-based sampling;
5. *Sampling rate (or sampling rate per campaign)*, which corresponds to the frequency of measurements;
6. *Sensor methods*, which corresponds to the automatic tools possibly used to collect measurements.

DSMM does not instead contain information explicitly referred to EBVs.

For this reason, we analysed the DSMM to identify elements suitable for the provision of the information suggested by Kissling et al. (2018a) to identify and aggregate data provided by any research infrastructure to build global data products, i.e.:

a. the EBVs measured, that can be found in the elements "Abstract", "Title" and "Keywords" of DSMM;
b. the EBVs dimensions (taxonomy, time and space);
c. the EBVs attributes for each dimension (extent, resolution and measurement units);
d. the EBVs uncertainties related to the measurement of each dimension;
e. the workflow steps accomplished to provide a usable-EBV dataset with reference to the standard data processing pipeline described in the above-mentioned work. The first three steps are: (1) identify and import raw data and associated metadata; (2) check data-sharing agreements and licences; (3) check data completeness and consistency. We considered the accessibility and completeness of metadata with respect to this information a requirement itself.

Table 1 reports the mapping of b. c. and d. to DSMM elements, while Table 2 details the mapping of the workflow steps described in e.

The mapping was obtained by analysing the model and selecting suitable elements to provide the information considered and by checking them with compiled metadata.

Collection of EBV information from metadata

The EBV information, described in the previous section, was collected for every site of LTER-Italy and structured in a database. The steps that we followed to collect the EBV information from the metadata elements, both of the sites and of the datasets, are presented below:

1. The investigator accesses the metadata through the public web interface of DEIMS-SDR and reads the content of the selected metadata elements exposed in a human-readable format. Through DEIMS-SDR, it is possible to read sites and datasets meta-

data shared by the network and, in particular, the values for the elements synthesised in Table 3. As reported in Kliment and Oggioni (2011), dataset metadata can be automatically exported according to standard XML schema Ecological Metadata Language (EML) and site metadata can be exported according to the standard schema Environmental Monitoring Facilities (EF). However, some information actually stored following DEIMS schemas, has no counterpart in these machine-readable standard schemas exposed by the platform. For instance, the "eLTER parameter" element of site metadata, which plays a key role in our analysis, is missing. For this reason, the process cannot be entirely automated and manual retrieval of values is required.

2. The investigator records the values of the variables under consideration for every site in a database. This database constituted the groundwork from which we derived the descriptive statistics presented in this study: it is publicly available in the form of a spreadsheet (Zilioli and Oggioni 2018).
3. The investigator uses the database to identify two lists of sites:
   a) the list of sites declaring SA and SD activities, obtained from site metadata;
   b) the list of sites exposing SA and SD related datasets, obtained from dataset metadata;

Table 1. Information suitable for building EBV data products mapped to DEIMS-SDR DSMM elements. The table illustrates elements which report information on EBV dimensions, attributes and uncertainties. The name of the related fields appears between parentheses while "ND" is used when elements to report the information are missing in the model.

EBV Dimensions ▶ **EBV Attributes and Uncertainties ▼**	**Taxonomy**	**Space**	**Time**
Extent			
(e.g. how many and which species are documented; sampling locations, satellites; length of time series, continuous recording, time period of collection of records)	Taxonomic coverage - Biological classification (field_bio_classification) OR Keywords (field_keywords_envthes) OR Title (field_title)	Geographic (field_related_sites)	Temporal extent (field_date_range)
Resolution			
(e.g. species, genus, higher taxonomic level; volume, resolution of satellite sensors; time window of sampling; sampling frequency)	Keywords (field_keywords_envthes) OR Title (field_title)	Abstract (field_abstract)	Sampling time span (field_sampling_time_span)
Measurement units			
(e.g. taxonomic entity for which species distribution and abundance data are sampled; metres, cubic metres, degrees; hours, days, months, years, decades)	Title (field_title) OR Abstract (field_abstract)	Method (field_related_links)	Minimum sampling unit (field_minimum_sampling_unit)
Uncertainties			
(e.g. wrongly recorded coordinates; precision of time of collection; identification and observation uncertainty differences in taxon concepts)	ND	ND	ND

Table 2. Information suitable for building EBV data products mapped to DEIMS-SDR DSMM elements. The table associates workflow steps required to build EBV-usable datasets to DSMM elements carrying the appropriate information. The name of the related field appears between parentheses.

Workflow Step ▼	DSMM elements ▼
Identify and import raw data and associated metadata (1)	Data set Title (field_title); DOI (field_doi), Online location (field_online_locator)
Check data-sharing agreements and licences (2)	Principal and granted permission (field_access_use_termref) Intellectual rights (field_dataset_rights)
Check data completeness and consistency (3)	Quality assurance (field_quality_assurance)

The total number of LTER-Italy sites of list a) are used to measure the EBVs Potential Contribution (PC) of the network; the total number of LTER-Italy sites of list b) are used to measure the Actual Contribution (AC) of the network.

Assessment of EBVs Potential Contribution (PC)

We measure the potential capacity of LTER-Italy to provide SA (or SD) data as the number of sites monitoring the selected variable against the total number of sites in our sample, as formalised in the following formula:

$$PC_v(LTER-Italy) = \frac{SM_v(LTER-Italy)}{S_{Total}},\ v \in \{SA, SD\} \subset EBV$$

where PC_v (LTER-Italy) is the Potential Contribution of LTER-Italy to EBV variable v. EBV in the formula represents the set of all EBVs: In our study, we are limiting v only to SA (Species Abundance) or SD (Species Distribution). SM_v is the number of Sites with the site metadata compiled for variable v. S_{total} is the number of sites taken into consideration as described in the "Case study" subsection.

Assessment of EBVs Actual Contribution (AC)

We measure the actual capacity of LTER-Italy to provide SA (or SD) usable data as the number of sites providing at least one dataset metadata compiled for the selected variable against the total number of sites in our sample (S_{total}):

$$AC_v(LTER-Italy) = \frac{SDM_v(LTER-Italy)}{S_{Total}},\ v \in \{SA, SD\} \subset EBV$$

where AC_v (LTER-Italy) is the Actual Contribution of LTER-Italy to EBV variable v. In our study, v is limited to SA (Species Abundance) or SD (Species Distribution) amongst

Table 3. Selection of Site Metadata and Data Set Metadata elements for analysis.

Site Metadata	Data Set Metadata
Site name	Dataset title
eLTER parameters (Biodiversity (EBV) – Object (taxon))	Related site
	Abstract
	Keywords
	Access and use constraints
	Intellectual rights
	Online distribution
	Geographic
	Temporal extent
	Taxonomic coverage

other possible EBVs. SDM_v is the number of Sites with at least one dataset metadata in which one of the elements reported in Table 3 is compiled for variable v. Again, S_{total} is the number of sites taken into consideration, as described in the "Case Study" subsection.

Collection and Assessment of Data management practices

For the considered LTER-Italy sites, we also imported the values of metadata elements belonging to the "Data management" and "Data sharing policies" sections, which contain additional information about data handling and sharing practices. We decided to enrich the EBV information retrieved through eLTER parameters with that describing the data management practices exposed in Site Metadata, so as to identify the researcher's attitude towards sharing data with external users. Although these data management practices are declared by the site managers in relation to their whole activity and not specifically referred to EBVs, we consider this information suitable for describing technological characteristics of the site (e.g. storage media and formats used, web services created and general policies applied to ecological data) and helpful to explain discrepancies between PC and AC.

For this assessment, we selected the following elements of SMM:

- *Data Storage Location*: this element describes the general design of data storage and the number of storage locations for data. By compiling this element, the site manager provides information on central or distributed data storage, the number of storage locations within an organisation as well as the storage locations situated by other organisations;
- *Storage format*: this element describes the different formats in which data are managed or are available to end users;
- *Data services*: this element describes which services are provided to end users (external or internal) to connect to data;
- *General data policy*: this element describes which rewarding actions and restrictions to user or activities are applied to data;
- *Data Request Format*: this element describes how the datasets can be requested from the site.

Results

The application of the methodology to the LTER-Italy case study resulted in the outcomes presented in this section.

EBVs Potential Contribution from LTER-Italy

The Potential Contributions from LTER-Italy are:

PC_{SA} (LTER – Italy) = 53%
PC_{SD} (LTER – Italy) = 42%

It is possible to group the sites in accordance to the biome they declare to monitor as explained in subsection "Mapping EBV information". Figures 3, 4 and 5 represent these groups of sites, the "Terrestrial", "Marine" and "River and Lakes", respectively. In these figures, we present not only the number of sites measuring SA and SD, but also the number of sites measuring other EBVs, to contextualise the contribution of the network to EBVs framework in a wider perspective.

The figures profile each biome-specific group with respect to the whole set of EBVs and each bar counts the number of sites declaring activities related to the corresponding EBV. Hence, through this analysis of metadata, we can compare our main analysed potential contributions (for SA and SD) with other EBVs monitored by LTER-Italy sites.

For marine biome, LTER-Italy accounts for 10 sites as potential providers for SA and SD EBV measures, which represent 23% of the sample in both cases. For terrestrial biome, LTER-Italy accounts for 8 and 7 sites for SA and SD EBV, respectively, which represent 19% and 16%; for the River and Lakes biome, the network accounts 5 (SA) and 1 (SD) sites, i.e. 12% and 2%, respectively.

SA and SD are the most measured EBVs. We can distinctly consider the total number of sites for every biome and restrict the analysis to them. In this case, the evaluation of the potential contribution to each biome is: 100% for SA, 88% for SD in Terrestrial biome; 77% for SA, 77% for SD in Marine biome; 50% for SA, 10% for SD in Rivers and Lakes biome.

Although there is a high number of biodiversity monitoring sites for both EBVs in each biome, the analysis suggests the presence of bias in long-term monitoring of biodiversity for EBVs, different from SA and SD. In fact, "Genetic composition" is an under-represented EBV class as only one site in LTER-Italy provides measures for "Allelic diversity" EBV; moreover, with respect to the six GEO BON classes that groups EBV (see "Introduction" section), while the Marine and Terrestrial biomes of Italy can be potentially described with 4 of the 5 class of EBVs, the Rivers and Lakes biome can be potentially described by data which cover only 2 out of 5 classes. By considering together SA and SD, monitoring sites which are potentially able to provide useful data are 72% of our sample.

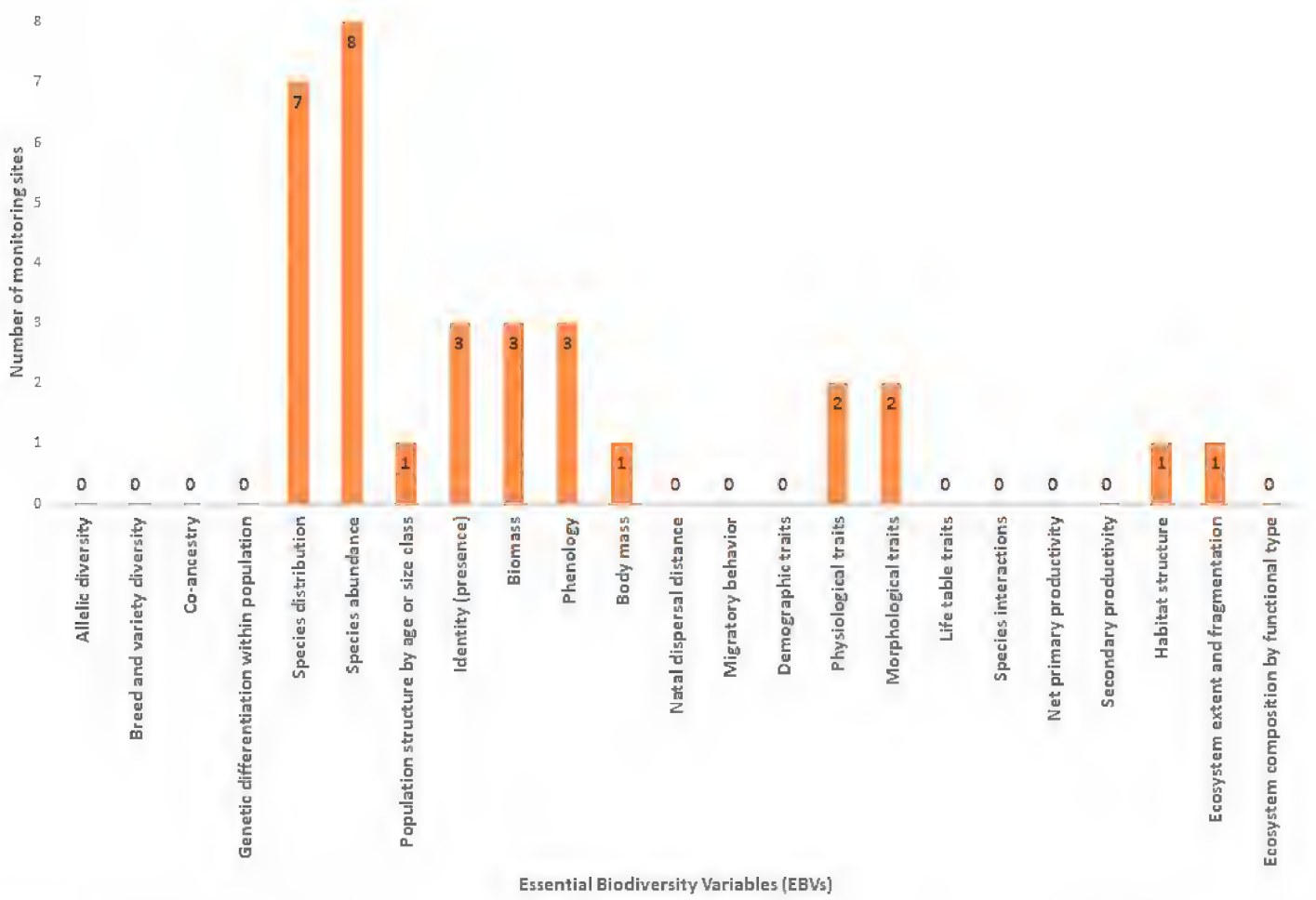

Figure 3. EBVs coverage for Terrestrial Biome sites of LTER-Italy (total number 8). The height of the bars represents the number of Terrestrial sites that are claimed to measure the selected variable.

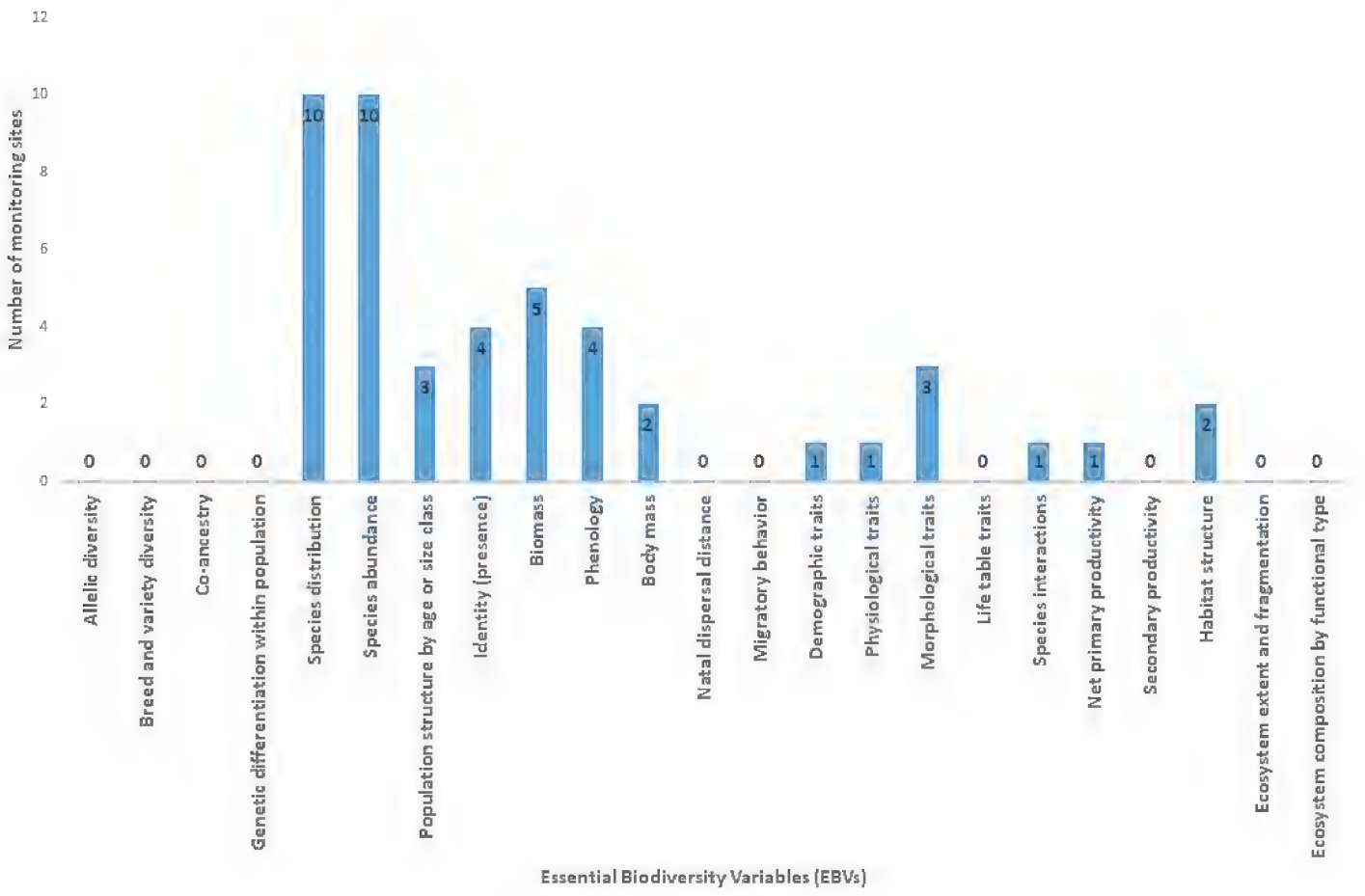

Figure 4. EBVs coverage for Marine Biome sites of LTER-Italy (total number 13). The height of the bars represents the number of Marine sites that are claimed to measure the selected variable.

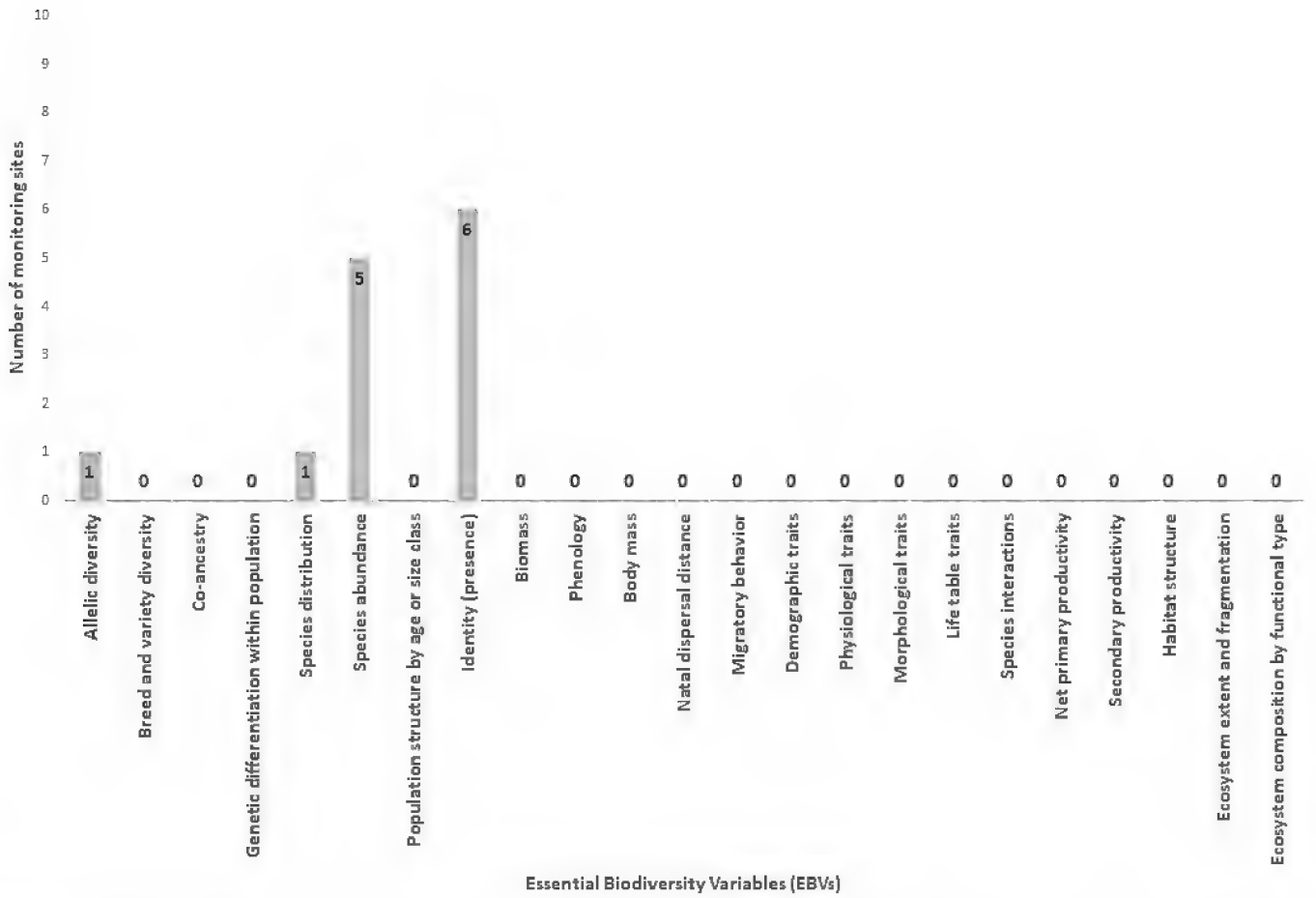

Figure 5. EBVs coverage for Rivers and Lakes Biome sites of LTER-Italy (total number 10). The height of the bars represents the number of Freshwaters sites that are claimed to measure the selected variable.

EBVs Actual Contribution from LTER-Italy

The Actual Contributions from LTER-Italy are:

AC_{SA} (LTER – Italy) = 14%
AC_{SD} (LTER – Italy) = 14%

The two contributions are the same because all the dataset metadata are referred to as "Species Abundance", thus providing measures of presence and absence of species which are useful also for "Species Distribution".

We can expand this numeric result with some further consideration on EBV dimensions (taxonomy, space, time) and attributes (extent, resolution, measurement unit, uncertainties), trying to evince the adequacy of metadata with respect to those discussed in section "Materials and Methods".

Figure 6 presents the percentage of completeness of dataset metadata with respect to these information requirements, by considering the full (100%) metadata completeness when information is given in each dimension of EBVs (taxonomy, space and time) for the three attributes (extent, resolution and measurement units).

Regarding the observed taxonomic groups, metadata are accessible for plankton (phytoplankton, zooplankton) and vascular plants. The orientation of the network to focus on these taxonomic groups is confirmed by the analysis of site metadata by

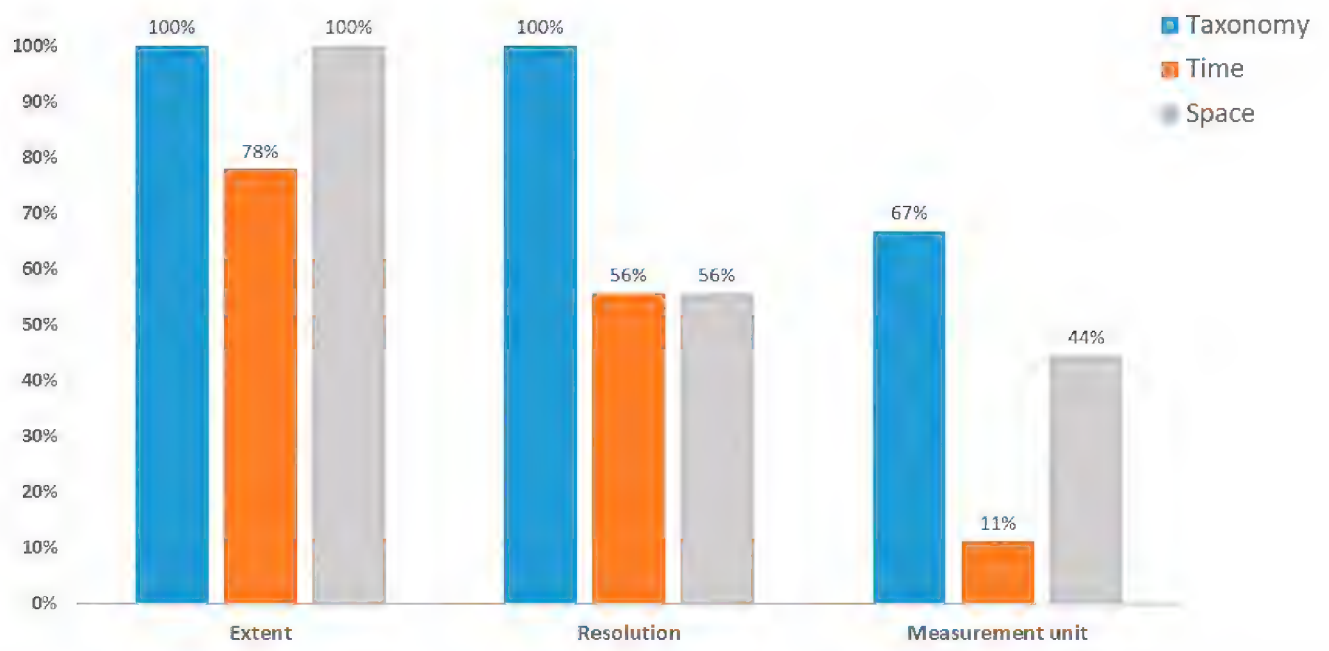

Figure 6. Metadata completeness with respect to EBV dimensions (taxonomy, space and time) and attributes (extent, resolution and measurement units) as mapped in Table 1.

which we observe that 29 sites can provide abundance measures for Phytoplankton and 13 sites for Zooplankton. Even if, in 100% of cases, the metadata element providing information for the taxonomy extent is compiled, the terms used do not belong to the ecological controlled vocabulary: Identification of organisms is given through free texts defining heterogeneous groups of taxonomic categories. Traditional methods (e.g. vegetation surveys, cells counting) are used to provide data along different spatial and temporal extents as described in "Materials and Methods".

Metadata are provided for long time-series datasets covering about 25–30 years or shorter periods. The 78% of metadata illustrated a sampling frequency time of five months, but resolution is provided by 56% and measurement units are not provided for 90% of metadata. In 100% of cases, sampling areas are carefully georeferenced through the metadata element "Geographic", reporting information about the spatial extent with altitudes and bounding coordinates provided by geotagging devices. However, also in this case, resolution and measurement units are provided in 56% and in 44% of metadata, respectively.

Figure 7 shows the completeness of metadata elements presented in Table 2, required by the workflow to integrate primary data in an EBV data product. A total of 100% of metadata and datasets is identified by elements about online distribution or by a title. In particular, the "Resource Locator" element accounts for the available ways to access the dataset resource: it is the navigation section of a metadata record, pointing users to the location where a dataset can be retrieved or where information about how to acquire a dataset can be obtained (e.g. the Uniform Resource Locator, URL; the E-mail to request a dataset). The "Resource Locator" element can also consist of a Digital Object Identifier (DOI) pointing directly and persistently to the dataset. This element covers two cases: The first is when an interested person must write asking for the data and, in this case, an e-mail address is provided; the second is when data are published

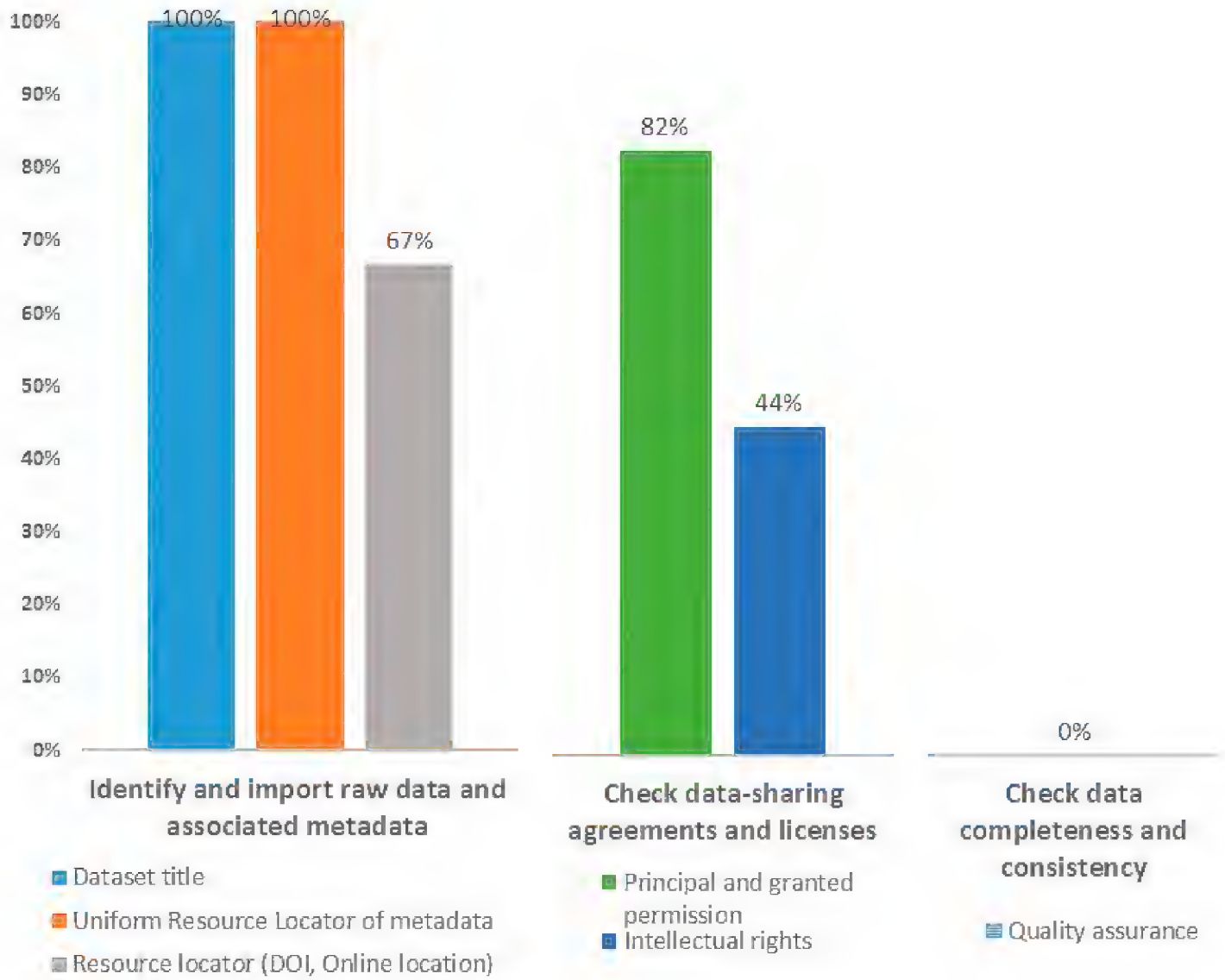

Figure 7. Metadata completeness with respect to the information mapped in Table 2. The three stepsrequired to obtaining EBV-usable datasets (i.e. datasets with measurements and observation protocols in thecorrect formats) are reported. Bars of different colours correspond to percentages of information provisionneeded for checking the workflow steps through different metadata elements. In particular, bar 1 (lightblue) corresponds to the completeness percentage for "Dataset Title"; bar 1 (orange) for "Uniform ResourceLocator of metadata"; bar 1 (grey) for "Resource Locator" of dataset; bar 2 (green) for "Principal and granted permissions"; bar 2 (blue) for "Intellectual rights"; bar 3 for "Quality assurance".

online and are retrievable through the provided URL or DOI. We highlighted in the figure the percentage of dataset URLs, 67%. The presence of URLs should indicate the online availability of a dataset stored within remote servers, project or institutional online repository, but unfortunately, none of the datasets proved to be publicly available, neither to be downloaded nor to be viewed, thus limiting access to the structure, the format and the observation protocols used to create data.

Licences and data-sharing agreements are applied to 82% of datasets through the metadata element "Principal and granted permission". In particular, there are distinctions in licensing based on intended use of datasets (for research, for public). For research uses, the actual granted permissions are "Free for access and use upon request" and "Free for access", while for generic public uses, the "Other restrictions according to rules defined in in-

tellectual rights" are applied by the providers and more finely defined by the metadata field "Intellectual rights". The field "Intellectual rights" is specified for 44% of datasets and, in the case of generic public uses, it almost always asks for "co-authorship on publications resulting from the use of dataset". We found just one dataset with "No access" granted.

Data quality information is not provided for any dataset; hence no dataset appears to be EBV-usable at metadata analysis level.

Data management practices

Figure 8 describes data management practices of potential EBVs contributors with respect to characteristics defined in subsection "Collection and Assessment of Data Management practices" of "Materials and Methods".

Data storage location is "central" (i.e. in the server of an institution) for 79% of sites, while in 10% of sites, data are distributed amongst repositories of different institutions and, in 11% of sites, data are distributed within the same institution (i.e. multiple places for data within the organisation that maintains and manages data).

With respect to *storage format*, 62% of the sites organise their data in structured files or spreadsheets, while 21% of sites declare their management of spatial datasets. Finally, dataset's proprietary formats are chosen by only 7% of sites.

Services for data access are not specified by 72% of sites while 14% exploit standard web services and 7% declare sharing its datasets through a generic "data portal".

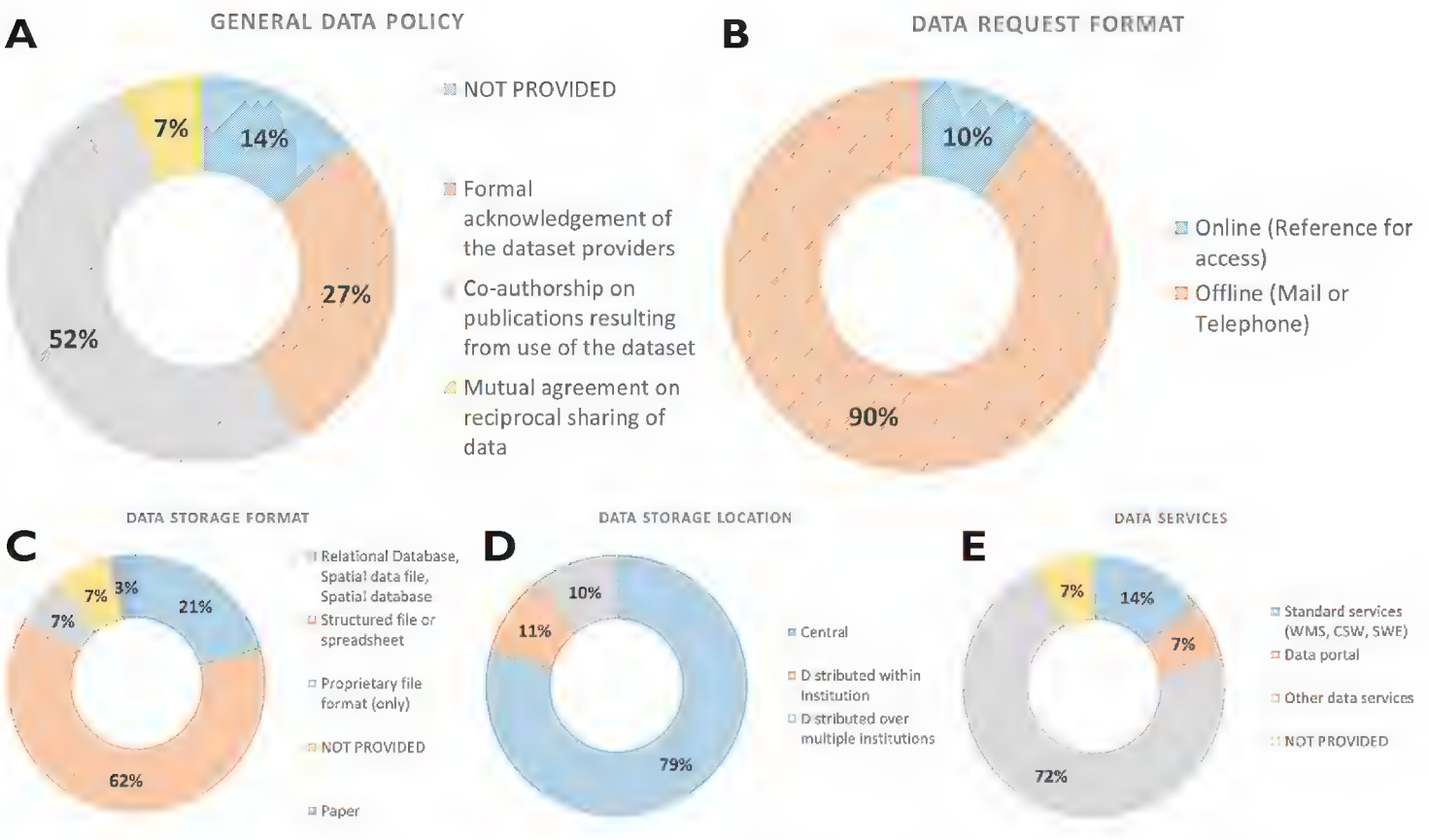

Figure 8. Data management practices associated with EBVs potential contributor sites. The Figure separately illustrates the relative percentages of sites for **A** policies applied to data **B** request formats for release data to external users **C** storage formats **D** storage location **E** web services used to make access to data.

A general preference for offline release of data, that explains the Actual Contribution results, is evident in the analysis of the *data request format*: only 10% of sites give online access to data, while 90% of sites prefer to be contacted by telephone or mail for giving access to data.

Finally, focussing on the *general data policy,* the data usage must be acknowledged by 52% of sites through demand for co-authorship on publications resulting from the use of datasets; mutual agreement on reciprocal data sharing are required to data users in 7% of cases only, while information is not provided at all by 14% of sites.

Discussion

Researchers and policy-makers are called to take joint actions to face biodiversity emergencies, as highlighted by the growing demand for readily accessible data that can be integrated and analysed in support of political decisions (Hardisty et al. 2013; Hoffmann et al. 2014). Even if biodiversity management literature reports advances with works relating EBVs to governmental policy (Turak et al. 2017a), the information pertinent to these essential measurements can be hidden to public users in the Web for different reasons, spanning from technical obstacles (e.g. limits to data discovery) to legal constraints (e.g. restrictions applied to sensitive data). In this paper, we investigate whether LTER-Italy provides measures for SA and SD by freeing the analysis from the need to directly access data. It is of interest to examine which specific motivations possibly hamper the public accessibility to data.

For the discussion of results, it is important to consider the following. First, LTER research is driven by specific scientific questions, posed by individual scientists or groups. These programmes are typically decentralised, rarely harmonised at global level and unevenly distributed geographically (Haase et al. 2018). Moreover, the selection of biotic and abiotic variables to monitor is at the sites' discretion, according to the available instrumentation and ecological focus. To monitor biodiversity is not a mandate of any central funding body or any coordinating scientific committee, hence not all the sites are expected to provide these measures. Second, communicating biodiversity change to wider audiences remains challenging, even if necessary to make biodiversity measurements into effective management actions (Turak et al. 2017b). Thus, a comprehensive, trustworthy and synoptic overview of monitoring and research capacity of scientific networks is needed.

Through the analysis of EBV information derived from metadata, we described the potential and actual contribution of LTER-Italy to provide EBV related datasets for collection and mobilisation of SA and SD measures (Hardisty et al. 2019a). We demonstrate LTER-Italy's good potential in providing EBVs, but also the discrepancy in data provision for SA and SD, which is graphically represented in Figure 9.

In fact, while 53% of sites potentially provide SA and 42% of sites SD data (Figure 9a), the number of sites which actually collect datasets metadata are 14% for both (Figure 9b); moreover, no dataset is accessible due to web resource localisation problems (e.g. URL pointing to no resource, dead link, broken link, location shifting), thus limiting web users in accessing the primary data.

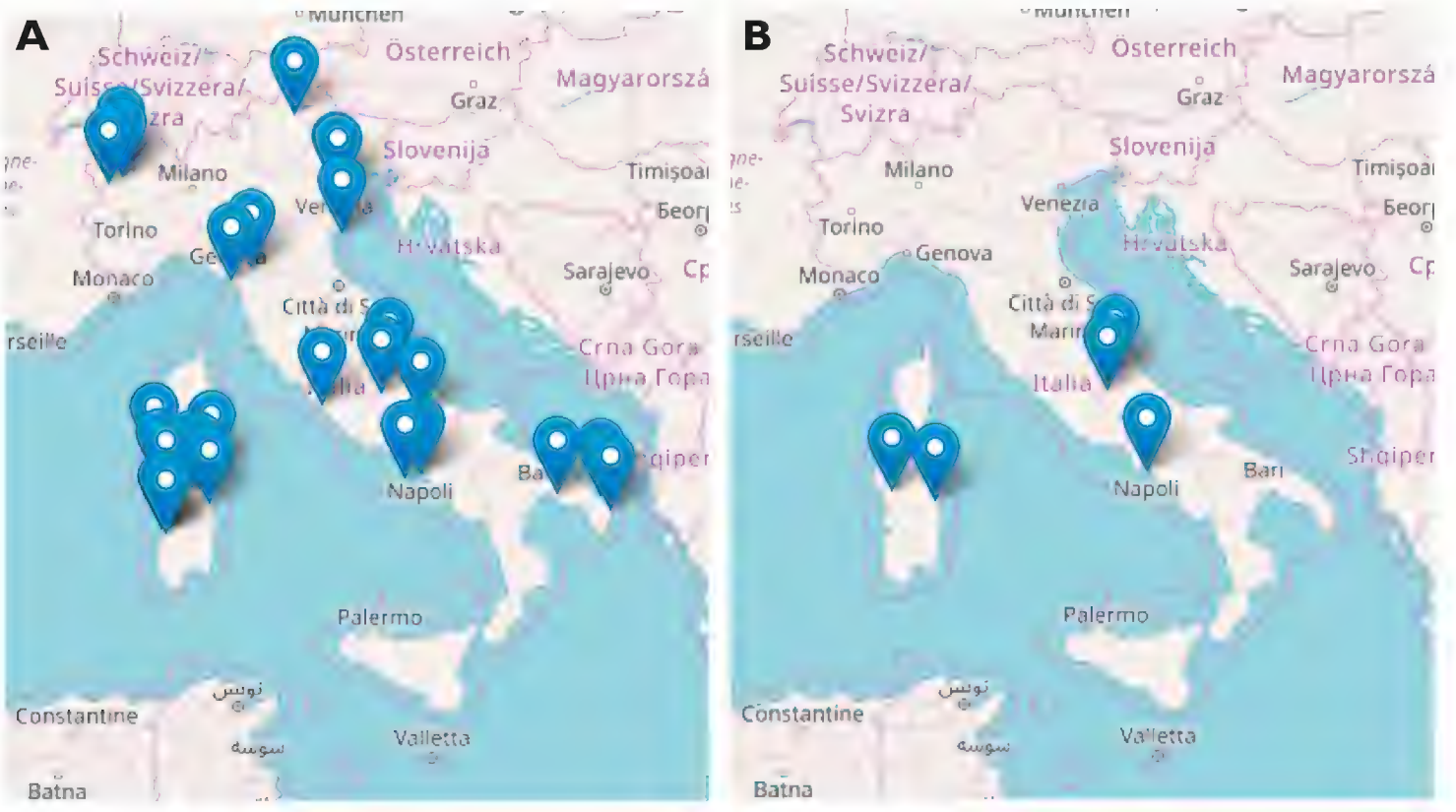

Figure 9. LTER-Italy potential and actual contribution sites. LTER-Italy sites which potentially supply SA and SD site-based, long-term measures are represented with a placeholder in **A** while sites which currently provide SA and SD metadata for primary datasets are represented in **B**.

Our metadata analysis suggests that community-related reasons are the factors which can explain the gap between the network's potential and actual capacity, thus providing clues to making data more accessible. Although several studies highlight that scientists often do not make their data available in digital form, for reasons including insufficient time and lack of funding (Tenopir et al. 2011), the analysis of data management practices rather suggests that the community is open to release its data, but preferably through offline media, instead of doing it by applying additional restriction to online distribution tools: this is consistent with our results that only a small part of the community (21% of sites) uses data-sharing services (standard Web services, data portals) and the greater part (79% of sites) centrally archives datasets, rather than distributing them through different storage media. However, these results are not conflicting with the more general attitude of the community to share data, but they pinpoint the need for tailored solutions to improve discoverability and reusability of data in this scenario. In fact, 52% of the general policies and 60% of licences, applied directly to biodiversity data, indicate that scientists approve data sharing for: (i) research purposes insofar as the collaboration is rewarded with citations or co-authorship (e.g. licences chosen are "Free for access and use upon request"; "Co-authorship on publications resulting from use of the dataset"); (ii) public purposes, insofar a formalised recognition, coming from the use of data, is given.

In such a context of limited online access to data, well-compiled metadata are even more necessary. Different types of metadata can compensate for the choice to regulate access to data, by supplying information for discovering and mining EBV information.

Different from the data management workflow described in Kissling et al. (2018a, 2018b), we found that published metadata of sites play an essential role in provid-

ing sound information about which EBVs are monitored. First, being compiled by site managers, they offer trustworthy indications about the observed properties of biodiversity. This means that the metadata contained in DEIMS-SDR are suitable wherever an authoritative assessment of measured parameters is needed, for example, when semi-quantitative or qualitative analysis are required for Ecosystem Integrity and Ecosystem Services assessments or for biodiversity change assessment (Turak et al. 2017a; Stoll et al. 2015), which are currently carried out with time-consuming surveys of key stakeholders and researchers. However, by providing the first mapping between EBVs metadata requirements and elements of DEIMS-SDR metadata models, we underline that technical improvements, facilitating the retrieval of EBVs information, have also to be addressed, particularly to assure its thorough exportation in standard formats (EML, EMF).

Second, metadata can be useful to identify thematic focus of any network (not only LTER) exposing metadata in DEIMS-SDR. In fact, through metadata analysis, we assess that LTER-Italy conducts biodiversity measures through different numbers of sites in every realm. Marine and terrestrial biomes are described with a higher number of EBV classes (5 and 4, respectively) with respect to freshwaters biome (2 classes) and with different frequencies for each EBV. SA and SD are the most measured EBVs, but the analysis shows that not all the sites provide these measures. The result can direct financial resources to activate monitoring activities, at least by volunteers of local communities through citizen science projects which present several advantages over traditional *in situ* field surveys for the collection of SA and SD data (Chandler et al. 2017, Kissling et al. 2018a).

The analysis of site metadata can provide spatial and temporal coverage, sampling frequency and monitored taxa, without the need for exploring related data, thus facilitating the planning of harmonised research activities at network scale. The method highlights, in fact, the capacity of the network in supplying data for taxa groups which are less monitored than invertebrates or vascular plants, towards which there is a bias described in the EBV-related literature (Proença et al. 2017). However, since metadata are compiled by site managers, they can be incomplete for elements that are not mandatory, as in the case of the "eLTER Parameter". For this specific reason, our analysis was limited to a sample of the community which is not representative for the comprehensive capacity of LTER-Italy to monitor EBVs. For example, ten sites are excluded from the analysis, as their metadata reported information on biodiversity solely in the element "Parameter". However, even if this element reports information related to generic species abundance and distribution measures, it provides information neither with respect to other measures referable to the EBV concept, nor at the level of detail required to obtain a comprehensive representation of the investigated objects and scales. To expand the statistics to the whole network, all sites have to describe research activities through specific metadata elements and the site metadata analysis needs to be completed with that related to datasets metadata. Through datasets metadata, we attempted to evaluate more deeply whether their information enables the reuse of datasets and whether datasets are accessible to other investigators: for example, to provide *in situ* data for Calibration and Validation activities of remote sensing analysis, as described in Mirtl et al. (2018). A dataset is deemed EBV-usable if (1) primary data and associated metadata are identified and imported, (2) data-sharing agreements and licences are checked and

(3) data completeness and consistency are described. Through the analysis of DEIMS-SDR, no dataset can be considered EBV-usable in the above sense. In fact, 67% of metadata provide an online location for data, partially satisfying (1); 82% of metadata satisfy (2), indicating data-sharing agreements associated to data and no metadata satisfy (3), offering information about quality check. This further enforces the need for metadata curation so as to assure visibility to EBVs monitoring activities of the network.

With respect to other worldwide providers, we conclude that LTER-Italy can contribute to SA and SD measures and that interoperability to integrate them with other data is partly achieved at two levels (Haslhofer and Klas 2010) as described below:

- Legal interoperability, which occurs at metadata level, where general data policies applied from sites, principal and granted permissions, as well as intellectual rights related to datasets, are specified.
- Technical interoperability, which occurs at metadata level and is assured by the DEIMS-SDR IT infrastructure, which allows the export of EBVs metadata in standard schema.

Nevertheless, these two levels are not fully achieved because (i) LTER-Italy dataset metadata just partially report how to allow the reuse of data without directly contacting owners and (ii) the implementation of mapping DEIMS-SDR metadata models to standard schemas needs to be completed. For these reasons, the next section is dedicated to suggestions for the improvement of both the IT infrastructure and the data provider support system, in order to expand the visibility of LTER sites with respect to SA and SD measures.

Conclusion and recommendations

The EBV concept should become the window into biodiversity observation systems upon which researchers, managers and decision-makers can better interact. Related web resources aid the streamlining of the EBV information exchange amongst different stakeholders insofar as its discovery and reuse are assured. The synoptic, comprehensive and harmonised overview of the set of local research which resulted by mining this information is of particular importance for LTER observational design purposes, as monitoring programmes need to be more coordinated and improved through sites' collaborations. This paper suggests a method, based on metadata analysis, to reveal capacities and gaps in these networks with disparate focuses on ecology to provide EBVs measurements. Since the present analysis exploits metadata of field observations, harmonised through the EBV concept and described in the DEIMS-SDR repository, it can be applied to every research organisation using this information system (e.g. Murgia Alta EcoPotential site does not belong to LTER-Italy, but its site managers can benefit from DEIMS-SDR metadata models to expose information), by offering an approach both to coordinate monitoring schemes for primary data collection and to evenly assess the role of Biodiversity Research Infrastructures (BRIs) (Hardisty et al. 2019a). LTER-Italy, the network to

which we apply this method, is a relevant case study, as it is deployed in a country that is extremely rich in biodiversity: it has the highest number and density of both animal and plant species within the European Union, as well as a high rate of endemisms (Convention on Biological Diversity, Country Profile, Italy). Thus, assessing the monitoring coverage of this system is essential for the conservation management of biological diversity and to centrally design its research activities, which actually represent a collection of individual monitoring studies that vary across time scales and research focuses.

Our results demonstrate a documented capacity to provide essential measures at two different levels of interoperability through the information system DEIMS-SDR, but underline the need to support the community and to optimise the EBV information retrieval to improve the assessment and hence the effectiveness of LTER as an observing system. The analysis behind this work also allows us to provide some recommendations regarding the tools proposed for the LTER network. As discussed, DEIMS-SDR can be exhaustively consulted only through the user interface and provides information on the attributes of each of the EBV dimension (see Table 1).

In order to provide the same analysis for different LTER networks or for a set of sites (e.g. those based on networks or projects), we would suggest:

1. to formally structure EBVs information both in SMM and in DSMM to (i) give visibility to those sites which choose to restrict the online data-sharing and (ii) to enable the automating of EBV information analysis through specific metadata elements. Particularly, we suggest:
   I. to complete the implementation of the mapping between DSMM and EML schema, following that which is described in Kliment and Oggioni (2011), where "Taxonomic coverage" within DMSS (field_bio_classification) was effectively mapped to the EML corresponding field (taxonomicClassification);
   II. to improve the description of datasets and their discovery: the DSMM should provide a field where the corresponding EBV or eLTER parameter could be inserted, as it currently happens for sites;
   III. to map the values of eLTER parameter field (field_elter_parameters) in the field observedProperty of the EF metadata exposed by DEIMS. Currently, through EF schema, only the contents of the "Parameter" field (field_parameters_taxonomy), i.e. description of the observed parameters and parameter groups at the site, are provided without the hierarchy of details for methods and instrumentations provided by eLTER parameters;
2. to ensure metadata completeness through curation staff to create a legacy of well-designed and documented long-term observations. In fact, the process of creating and publishing metadata is relatively new amongst scientists despite its value in domains like ecology, where metadata improve the reusability of data. For example, protocols and instruments information are needed to assure interpretation of data over time and to allow comparisons when different methods are adopted. Metadata compilation is error prone (Kervin et al. 2013) and also perceived as a burden from researchers which often results in incomplete metadata provision. Nevertheless, curation staff (e.g. data stewards, librarians, help desk) can support scientists by stimulating their willingness

to share (meta)data by identifying contextual causes which hamper the practice (Zilioli et al. 2019) or by lightening the compilation procedures with informatics facilities (Fugazza et al. 2016; Fugazza et al. 2018; Pavesi et al. 2016). Specifically addressing the issue of this paper, we suggest ensuring a careful recording of the EBVs information by employing dedicated personnel to assist scientists in creating reliable metadata.

Acknowledgements

The activities described in this paper have been partially funded by the Italian Flagship Project RITMARE, eLTER H2020 project, NextData, e-biodiversity RI Lifewatch Italy.

References

Barnosky AD, Matzke N, Tomiya S, Wogan GOU, Swartz B, Quental TB, Marshall C, McGuire JL, Lindsey EL, Maguire KC, Mersey B, Ferrer EA (2011) Has the Earth's sixth mass extinction already arrived? Nature 471(7336): 51–57. https://doi.org/10.1038/nature09678

Bubb P, Chenery A (2011) National Indicators, Monitoring and Reporting for the Strategic Plan for Biodiversity 2011–2020. UNEP-WCMC, Cambridge. https://www.bipindicators.net/system/resources/files/000/000/401/original/815.pdf?1480337853

Chandler M, See L, Copas K, Bonde AMZ, López BC, Danielsen F, Legind JK, Masinde S, Miller-Rushing AJ, Newman G, Rosemartin A, Turak E (2017) Contribution of citizen science towards international biodiversity monitoring. Biological Conservation 213: 280–294. https://doi.org/10.1016/j.biocon.2016.09.004

DEIMS-SDRMetadata Models (2018) https://deims.org/models

Díaz S, Fargione J, Chapin FS, Tilman D (2006) Biodiversity Loss Threatens Human Well-Being. PLoS Biology 4(8): e277. https://doi.org/10.1371/journal.pbio.0040277

Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJB, Collen B (2014) Defaunation in the Anthropocene. Science 345(6195): 401–406. https://doi.org/10.1126/science.1251817

eBird (2018) https://ebird.org/home

European Biodiversity Portals, EBP (2018) http://biodiversity.eubon.eu/

Fugazza C, Pepe M, Oggioni A, Tagliolato P, Pavesi F, Carrara P (2016) Describing Geospatial Assets in the Web of Data: A Metadata Management Scenario. ISPRS International Journal of Geo-Information 5(12): 229. https://doi.org/10.3390/ijgi5120229

Fugazza C, Pepe M, Oggioni A, Tagliolato P, Carrara P (2018) Raising Semantics-Awareness in Geospatial Metadata Management. ISPRS International Journal of Geo-Information 7(9): 370. https://doi.org/10.3390/ijgi7090370

Geijzendorffer IR, Regan EC, Pereira HM, Brotons L, Brummitt N, Gavish Y, Haase P, Martin CS, Mihoub J-B, Secades C, Schmeller DS, Stoll S, Wetzel FT, Walters M (2016) Bridging the gap between biodiversity data and policy reporting needs: An Essential Biodiversity Variables perspective Cadotte M (Ed.) Journal of Applied Ecology 53: 1341–1350. https://doi.org/10.1111/1365-2664.12417

GenBank (2018) https://www.ncbi.nlm.nih.gov/genbank/

GEO BON (2018) https://geobon.org/ebvs/what-are-ebvs/

GLOBIS-B (2018) http://www.globis-b.eu/news/11/globis-b_contributes_to_a_better_definition_of_ebvs.html

Global Biodiversity Information Facility, GBIF (2018) https://www.gbif.org/

Haase P, Tonkin JD, Stoll S, Burkhard B, Frenzel M, Geijzendorffer IR, Häuser C, Klotz S, Kühn I, McDowell WH, Mirtl M, Müller F, Musche M, Penner J, Zacharias S, Schmeller DS (2018) The next generation of site-based long-term ecological monitoring: Linking essential biodiversity variables and ecosystem integrity. The Science of the Total Environment 613–614: 1376–1384. https://doi.org/10.1016/j.scitotenv.2017.08.111

Hardisty A, Roberts D, The Biodiversity Informatics Community (2013) A decadal view of biodiversity informatics: Challenges and priorities. BMC Ecology 13(1): 16. https://doi.org/10.1186/1472-6785-13-16

Hardisty A, Michener WK, Agosti D, García EA, Bastin L, Belbin L, Bowser A, Buttigieg PL, Canhos DAL, Egloff W, De Giovanni R, Figueira R, Groom Q, Guralnick RP, Hobern D, Hugo W, Koureas D, Ji L, Los W, Manuel J, Manset D, Poelen J, Saarenmaa H, Schigel D, Uhlir PF, Kissling DW (2019a) The Bari Manifesto: An interoperability framework for essential biodiversity variables. Ecological Informatics 49: 22–31. https://doi.org/10.1016/j.ecoinf.2018.11.003

Hardisty AR, Belbin L, Hobern D, McGeoch MA, Pirzl R, Williams KJ, Kissling WD (2019b) Towards Essential Biodiversity Variables data products for monitoring alien invasive species. Environmental Research Letters 14: 2. https://doi.org/10.1088/1748-9326/aaf5db

Haslhofer B, Klas W (2010) A survey of techniques for achieving metadata interoperability. ACM Computing Surveys 42(2): 1–37. https://doi.org/10.1145/1667062.1667064

Hoffmann A, Penner J, Vohland K, Cramer W, Doubleday R, Henle K, Kõljalg U, Kühn I, Kunin W, Negro JJ, Penev L, Rodríguez C, Saarenmaa H, Schmeller D, Stoev P, Sutherland W, Ó Tuama É, Wetzel F, Häuser CL (2014) The need for an integrated biodiversity policy support process – Building the European contribution to a global Biodiversity Observation Network (EU BON). Nature Conservation 6: 49–65. https://doi.org/10.3897/natureconservation.6.6498

INSPIRE Thematic Working Group "Environmental Monitoring Facilities" (2012) D2.8.III.7 INSPIRE Data Specification on Environmental Monitoring Facilities – Draft Guidelines. European Commission Joint Research Centre. INSPIRE Implementing Rules for Metadata. Available from: https://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_EF_v3.0.pdf

Convention on Biological Diversity, Country Profile, Italy (2018) https://www.cbd.int/doc/world/it/it-nr-05-en.pdf

Kelling S, Hochachka WM, Fink D, Riedewald M, Caruana R, Ballard G, Hooker G (2009) Data-intensive Science: A New Paradigm for Biodiversity Studies. Bioscience 59(7): 613–620. https://doi.org/10.1525/bio.2009.59.7.12

Kervin KE, Michener WK, Cook RB (2013) Common Errors in Ecological Data Sharing. Journal of Escience Librarianship 2(2). https://doi.org/10.7191/jeslib.2013.1024

Kissling WD, Ahumada JA, Bowser A, Fernandez M, Fernández N, García EA, Guralnick RP, Isaac NJB, Kelling S, Los W, McRae L, Mihoub J-B, Obst M, Santamaria M, Skidmore AK, Williams KJ, Agosti D, Amariles D, Arvanitidis C, Bastin L, De Leo F, Egloff W, Elith J, Hobern D, Martin D, Pereira HM, Pesole G, Peterseil J, Saarenmaa H, Schigel D,

Schmeller DS, Segata N, Turak E, Uhlir PF, Wee B, Hardisty AR (2018a) Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale: Building global EBVs. Biological Reviews of the Cambridge Philosophical Society 93(1): 600–625. https://doi.org/10.1111/brv.12359

Kissling WD, Walls R, Bowser A, Jones MO, Kattge J, Agosti D, Amengual J, Basset A, van Bodegom PM, Cornelissen JHC, Denny EG, Deudero S, Egloff W, Elmendorf SC, Alonso García E, Jones KD, Jones OR, Lavorel S, Lear D, Navarro LM, Pawar S, Pirzl R, Rüger N, Sal S, Salguero-Gómez R, Schigel D, Schulz K-S, Skidmore A, Guralnick RP (2018b) Towards global data products of Essential Biodiversity Variables on species traits. Nature Ecology & Evolution 2(10): 1531–1540. https://doi.org/10.1038/s41559-018-0667-3

Kliment T, Oggioni A (2011) A1.2.2b - EnvEurope (LTER-Europe) Metadata Specification for Dataset Level. 16pp

LifeWatch (2018) https://www.lifewatch.eu/home

LPI (2018) http://www.livingplanetindex.org/about

LTER-Europe (2018) http://www.lter-europe.net

Michener WK (2006) Meta-information concepts for ecological data management. Ecological Informatics 1(1): 3–7. https://doi.org/10.1016/j.ecoinf.2005.08.004

Michener WK, Brunt JW, Helly JJ, Kirchner TB, Stafford SG (1997) Nongeospatial Metadata for the ecological sciences. Ecological Applications 7: 330–342. https://doi.org/10.1890/1051 0761(1997)007[0330:NMFTES]2.0.CO;2

Mirtl M, Borer T, Djukic E, Forsius I, Haubold M, Hugo H, Jourdan W, Lindenmayer J, Mc D, Dowell WH, Muraoka H, Orenstein DE, Pauw JC, Peterseil J, Shibata H, Wohner C, Yu X, Haase P (2018) Genesis, goals and achievements of Long-Term Ecological Research at the global scale: A critical review of ILTER and future directions. The Science of the Total Environment 626: 1439–1462. https://doi.org/10.1016/j.scitotenv.2017.12.001

Mollenhauer H, Kasner M, Haase P, Peterseil J, Wohner C, Frenzel M, Mirtl M, Schima R, Bumberger J, Zacharias S (2018) Long-term environmental monitoring infrastructures in Europe: Observations, measurements, scales, and socio-ecological representativeness. The Science of the Total Environment 624: 968–978. https://doi.org/10.1016/j.scitotenv.2017.12.095

Muller-Karger FE, Miloslavich P, Bax NJ, Simmons S, Costello MJ, Sousa Pinto I, Canonico G, Turner W, Gill M, Montes E, Best BD, Pearlman J, Halpin P, Dunn D, Benson A, Martin CS, Weatherdon LV, Appeltans W, Provoost P, Klein E, Kelble CR, Miller RJ, Chavez FP, Iken K, Chiba S, Obura D, Navarro LM, Pereira HM, Allain V, Batten S, Benedetti-Checchi L, Duffy JE, Kudela RM, Rebelo L-M, Shin Y, Geller G (2018) Advancing Marine Biological Observations and Data Requirements of the Complementary Essential Ocean Variables (EOVs) and Essential Biodiversity Variables (EBVs) Frameworks. Frontiers in Marine Science 5: 211. https://doi.org/10.3389/fmars.2018.00211

Murgia Alta EcoPotential site (2018) https://deims.org/bbd428c6-c4b8-4f47-9298-2b969d38a664

Pangaea (2018) https://www.pangaea.de/

Pavesi F, Basoni A, Fugazza C, Menegon S, Oggioni A, Pepe M, Tagliolato P, Carrara P (2016) EDI – A Template-Driven Metadata Editor for Research Data Journal of Open Research Software, 4: e40. https://doi.org/10.5334/jors.106

Pereira HM, Ferrier S, Walters M, Geller GN, Jongman RHG, Scholes RJ, Bruford MW, Brummitt N, Butchart SHM, Cardoso AC, Coops NC, Dulloo E, Faith DP, Freyhof J,

Gregory RD, Heip C, Hoft R, Hurtt G, Jetz W, Karp DS, McGeoch MA, Obura D, Onoda Y, Pettorelli N, Reyers B, Sayre R, Scharlemann JPW, Stuart SN, Turak E, Walpole M, Wegmann M (2013) Essential Biodiversity Variables. Science 339(6117): 277–278. https://doi.org/10.1126/science.1229931

Peterson GD, Harmàčkova ZV, Meacham M, Queiroz C, Jimènez-Aceituno A, Kuiper JJ, Malmborg K, Sitas N, Bennett EM (2018) Welcoming different perspectives in IPBES: "Nature's contributions to people" and "Ecosystem services.". Ecology and Society 23(1): art39. https://doi.org/10.5751/ES-10134-230139

Proença V, Martin LJ, Pereira HM, Fernandez M, McRae L, Belnap J, Böhm M, Brummitt N, García-Moreno J, Gregory RD, Honrado JP, Jürgens N, Opige M, Schmeller DS, Tiago P, van Swaay CAM (2017) Global biodiversity monitoring: From data sources to Essential Biodiversity Variables. Biological Conservation 213: 256–263. https://doi.org/10.1016/j.biocon.2016.07.014

Rosati I, Bergami C, Stanca E, Roselli L, Tagliolato P, Oggioni A, Fiore N, Pugnetti A, Zingone A, Boggero A, Basset A (2017) A thesaurus for phytoplankton trait-based approaches: Development and applicability. Ecological Informatics 42: 129–138. https://doi.org/10.1016/j.ecoinf.2017.10.014

Schmeller DS, Mihoub J-B, Bowser A, Arvanitidis C, Costello MJ, Fernandez M, Geller GN, Hobern D, Kissling WD, Regan E, Saarenmaa H, Turak E, Isaac NJB (2017) An operational definition of essential biodiversity variables. Biodiversity and Conservation 26(12): 2967–2972. https://doi.org/10.1007/s10531-017-1386-9

Stoll S, Frenzel M, Burkhardd B, Adamescuf M, Augustaitis A, Baeßler C, Bonet FJ, Carranzai ML, Cazacuf C, Cosor GL, Díaz-Delgadoj R, Grandink U, Haase P, Hämäläinenl H, Loke R, Müller J, Staniscii A, Staszewski T, Müller F (2015) Assessment of ecosystem integrity and service gradients across Europe using the LTER Europe network. Ecological Modelling 295: 75-87. https://doi.org/10.1016/j.ecolmodel.2014.06.019

Sutherland WJ, Bellingan L, Bellingham JR, Blackstock JJ, Bloomfield RM, Bravo M, Cadman VM, Cleevely DD, Clements A, Cohen AS, Cope DR, Daemmrich AA, Devecchi C, Anadon LD, Denegri S, Doubleday R, Dusic NR, Evans RJ, Feng WY, Godfray HCJ, Harris P, Hartley SE, Hester AJ, Holmes J, Hughes A, Hulme M, Irwin C, Jennings RC, Kass GS, Littlejohns P, Marteau TM, McKee G, Millstone EP, Nuttall WJ, Owens S, Parker MM, Pearson S, Petts J, Ploszek R, Pullin AS, Reid G, Richards KS, Robinson JG, Shaxson L, Sierra L, Smith BG, Spiegelhalter DJ, Stilgoe J, Stirling A, Tyler CP, Winickoff DE, Zimmern RL (2012) A Collaboratively-Derived Science-Policy Research Agenda von Elm E (Ed.). PLoS ONE 7: e31824. https://doi.org/10.1371/journal.pone.0031824

TEAM (2018) https://www.ncbi.nlm.nih.gov/pubmed/27665451

Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, Read E, Manoff M, Frame M (2011) Data Sharing by Scientists: Practices and Perceptions Neylon C (Ed.). PLoS ONE 6: e21101. https://doi.org/10.1371/journal.pone.0021101

Turak E, Brazill-Boast J, Cooney T, Drielsma M, DelaCruz J, Dunkerley G, Fernandez M, Ferrier S, Gill M, Jones H, Koen T, Leys J, McGeoch M, Mihoub J-B, Scanes P, Schmeller D, Williams K (2017a) Using the essential biodiversity variables framework to measure

biodiversity change at national scale. Biological Conservation 213: 264–271. https://doi.org/10.1016/j.biocon.2016.08.019

Turak E, Regan E, Costello MJ (2017b) Measuring and reporting biodiversity change. Biological Conservation 213: 249–251. https://doi.org/10.1016/j.biocon.2017.03.013

UNEP Convention on Biological Diversity (2010) UNEP/CBD/COP/DEC/X/2 Decision adopted by the Conference of the Parties to the Convention on Biological Diversity at its tenth meeting. Available from: https://www.cbd.int/doc/decisions/cop-10/cop-10-dec-02-en.pdf

UNEP Convention on Biological Diversity (2016) Remote Sensing of Essential Biodiversity Variables. UNEP Convention on Biological Diversity, subsidiary Body on Implementation, 2016 Available from: https://www.cbd.int/doc/meetings/sbi/sbi-01/information/sbi-01-inf-49-en.pdf

Vihervaara P, Auvinen A-P, Mononen L, Törmä M, Ahlroth P, Anttila S, Böttcher K, Forsius M, Heino J, Heliölä J, Koskelainen M, Kuussaari M, Meissner K, Ojala O, Tuominen S, Viitasalo M, Virkkala R (2017) How Essential Biodiversity Variables and remote sensing can help national biodiversity monitoring. Global Ecology and Conservation 10: 43–59. https://doi.org/10.1016/j.gecco.2017.01.007

Weltzin JF, Jones KD, Brown JF, Elmendorf S, Enquist C, Rosemartin A, Thorpe A, Wee B (2014) Using Essential Biodiversity Variables (EBVs) As a Framework for Coordination Between Research and Monitoring Networks: A Case Study with Phenology. AGU Fall Meeting Abstracts: GC51B-0404

Wilkinson MD, Dumontier M, Aalbersberg Ij J, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PA, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3: 160018. https://doi.org/10.1038/sdata.2016.18

Wohner C, Peterseil J, Oggioni A, Kliment T, Poursanidis D, Provenzale A (2018) Providing interoperable information for science: utilising INSPIRE to mobilise data from long - term observation sites and protected areas. In: Antwerp, Belgium. Available from: https://inspire.ec.europa.eu/events/conferences/inspire_2018/submissions/258.html

Wohner C, Peterseil J, Poursanidis D, Kliment T, Wilson M, Mirtl M, Chrysoulakis N (2019) DEIMS-SDR – A web portal to document research sites and their associated data. Ecological Informatics 51: 15–24. https://doi.org/10.1016/j.ecoinf.2019.01.005

Zilioli M, Oggioni A (2018) LTER-Italy EBVs and EI indicators inventory (Version 1.0) [Data set]. Zenodo. http://doi.org/10.5281/zenodo.1293554

Zilioli M, Lanucara S, Oggioni A, Fugazza C, Carrara P (2019) Fostering Data Sharing in Multidisciplinary Research Communities: A case study in the geospatial domain. Data Science Journal 18: 1. [in press]